Sparsity-Promoting Extended Kalman Filtering for Target Tracking in Wireless Sensor Networks

Engin Masazade, *Member, IEEE*, Makan Fardad, *Member, IEEE*, and Pramod K. Varshney, *Fellow, IEEE*

***Abstract*—In this letter, we study the problem of target tracking based on energy readings of sensors. We minimize the estimation error by using an extended Kalman filter (EKF). The Kalman gain matrix is obtained as the solution to an optimization problem in which a sparsity-promoting penalty function is added to the objective. The added term penalizes the number of nonzero columns of the Kalman gain matrix, which corresponds to the number of active sensors. By using a sparse Kalman gain matrix only a few sensors send their measurements to the fusion center, thereby saving energy. Simulation results show that an EKF with a sparse Kalman gain matrix can achieve tracking performance that is very close to that of the classical EKF, where all sensors transmit to the fusion center.**

***Index Terms*—Alternating directions method of multipliers, extended Kalman filter, sensor selection, sparsity-promoting optimization, target tracking, wireless sensor networks.**

I. Introduction

IN this letter, we study the target tracking problem in wireless sensor networks, where the aim is to estimate the position and velocity of an object emitting energy. In this work, we assume that each sensor measures the energy from the source, which attenuates as a function of the distance between the source and sensor locations. The fusion center collects the sensor measurements and is responsible for the final inference [1]. In WSNs, due to communications and energy constraints, it is often necessary to use limited number of sensors at each time instant of tracking. In the signal processing literature, there exist many sensor selection algorithms (see [2] and references therein). Information-based measures choose the sensor selection strategy that maximizes the expected gain in information [3], [4]. Sensor selection strategies that minimize the error in estimation are computationally more efficient than the information-based sensor selection methods [5] and have also been used in [2], [6]. Since selecting k informative sensors out of N sensors in the network is an NP-hard combinatorial problem, the sensor selection problem is relaxed and solved using convex optimization in [2]. Further, the combinatorial sensor selection problem is solved using compressed sensing algorithms in [7].

The extended Kalman filter (EKF) first linearizes the nonlinear measurement model, i.e., isotropic signal emission model, and then obtains the near-optimal Kalman gain matrix where the trace of the error covariance matrix is the objective function to be minimized. In this letter, we propose to obtain the Kalman gain matrix as the solution to an optimization problem in which a sparsity-promoting penalty function is added to the objective. The added term penalizes the number of nonzero columns of the Kalman gain matrix, which corresponds to the number of active sensors. Then, the resulting optimization problem is solved by using the alternating directions method of multipliers method (ADMM) [8]. The approach we take in this work for finding the optimal sparse Kalman filter gain closely follows the approach taken in [9], [10], which consider the problem of finding optimal sparse static feedback gains using ADMM. The results in [9], [10], [11], [13], demonstrate the effectiveness of the sparsity- promoting framework and the application of ADMM in rendering sparse matrix gains. In a sense, the present work can be considered as a dual to [9], [10]. ADMM has also been used for optimal sparse leader selection in networks of dynamical systems in [12].

The earlier works in sensor selection presented in [2], [4]–[7] assume that the number of selected sensors among all the sensors in the WSN at each time step of tracking is fixed and known. In this work, we do not enforce a specific number of active sensors. Instead, the number of selected sensors is determined by a parameter which characterizes the emphasis on the column sparsity of the Kalman gain matrix. Sparsity has been recently considered in target tracking applications [14], where the authors consider grid points at known positions in a region of interest (ROI) where the target could be potentially located. The grid points are the elements of the vector to be estimated, where only few of the grid points are nonzero by considering an isotropic signal emission model for the target. A sparsity-aware Kalman filter is then proposed to estimate values of the grid points where all the sensors transmit their measurements to the fusion center. On the contrary, in this work we do not assume that the target must be located at a grid point. Moreover, we assume that only a subset of sensors report to the fusion center at each time step.

The rest of the letter is organized as follows. Section II presents the system model for target tracking and the EKF process, In Section III we introduce the sparsity promoting Kalman filtering problem. Then, in Section IV we provide numerical examples and finally Section V concludes the letter.

Manuscript received July 03, 2012; revised August 21, 2012; accepted August 30, 2012. Date of publication September 21, 2012; date of current version October 16, 2012. The work of E. Masazade and P. K. Varshney was supported by the U.S. Air Force Office of Scientific Research (AFOSR) under Grant FA9550-10-1-0263. The work of M. Fardad was supported from National Science Foundation under award CMMI-0927509. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Shahrokh Valaee.

E. Masazade was with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA. He is now with the Department of Electrical and Electronics Engineering, Yeditepe University, Istanbul, Turkey (e-mail: engin.masazade@yeditepe.edu.tr).

M. Fardad and P. K. Varshney are with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: makan@syr.edu; varshney@syr.edu)

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2012.2220350

II. System Model

In this letter, the problem we seek to solve is to track a moving target using a WSN where N sensors are uniformly deployed in a square ROI of size b^2. Note that the uniform layout assumption is not necessary and target tracking based on sensor readings can be performed for an arbitrary network layout as long as sensor placements are known in advance. We assume that the target (e.g., an acoustic or an electromagnetic source) emits a signal from the location (x_t, y_t) at every discrete time instant t. All the sensors report their measurements to a central fusion center, which estimates the target state, i.e., the position (x_t, y_t) and the velocity of the target in the horizontal and the vertical directions $(\dot{x}_t, \dot{y}_t)$.

At time t, the target dynamics are defined by a 4×1-dimensional state vector $\mathbf{x}_t = [x_t \quad y_t \quad \dot{x}_t \quad \dot{y}_t]^T$. Target motion is defined by the following white noise acceleration model:

$$\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \boldsymbol{v}_t \tag{1}$$

where $\mathbf{F}$ is a 4×4 matrix which models the state dynamics and $\boldsymbol{v}_t$ is the process noise which is assumed to be white, zero-mean and Gaussian with the following covariance matrix $\mathbf{Q}$ which also has size 4×4.

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta & 0 \\ 0 & 1 & 0 & \Delta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{Q} = \tau \begin{bmatrix} \frac{\Delta^3}{3} & 0 & \frac{\Delta^2}{2} & 0 \\ 0 & \frac{\Delta^3}{3} & 0 & \frac{\Delta^2}{2} \\ \frac{\Delta^2}{2} & 0 & \Delta & 0 \\ 0 & \frac{\Delta^2}{2} & 0 & \Delta \end{bmatrix}. \tag{2}$$

In (2), Δ is the time interval between consecutive sensor measurements and τ is the process noise parameter. It is assumed that the fusion center has perfect information about the target state-space model (1) as well as the process noise statistics (2).

At any given time t, let the measurement vector be $\mathbf{z}_t \triangleq [z_{1,t}, \ldots, z_{N,t}]^T$, which has the form,

$$\mathbf{z}_t = h(\mathbf{x}_t) + \mathbf{w}_t \tag{3}$$

where $\mathbf{w}_t$ is the measurement noise representing the cumulative effects of sensor background noise and the modeling error of signal parameters. $\mathbf{w}_t$ is assumed to be white, zero-mean and Gaussian with the covariance matrix $\mathbf{R} = \sigma^2 \mathbf{I}_{N \times N}$, where $\mathbf{I}_{N \times N}$ is an identity matrix of size $N \times N$. The target is assumed to be an acoustic or an electromagnetic source that follows the isotropic attenuation model characterized by the function $h(.)$. As in [15], each element of $\mathbf{z}_t$, $z_{i,t}$, is the noisy signal amplitude received at the sensor i and is assumed to be expressible as,

$$z_{i,t} = \sqrt{\frac{P_0}{1 + (d_{i,t})^n}} + w_{i,t}. \tag{4}$$

In (4), P_0 denotes the signal power of the source, n is the signal decay exponent. $d_{i,t}$ is the distance between the target and the i^{th} sensor, $d_{i,t} = \sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}$, where (x_i, y_i) are the coordinates of the i^{th} sensor. Without loss of generality, n is assumed to be 2.

The extended Kalman filter (EKF) is the nonlinear version of the Kalman filter which linearizes the non-linear measurement model around the predicted state of the estimate [16]. The two processes of the EKF are briefly summarized as follows: In the prediction process, $\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}\hat{\mathbf{x}}_{t-1|t-1}$ is the predicted state estimate and $\mathbf{P}_{t|t-1} = \mathbf{F}\mathbf{P}_{t-1|t-1}\mathbf{F}^T + \mathbf{Q}$ is the covariance matrix of the predicted estimate. In the update process, the nonlinear measurement model is first normalized as $\mathbf{H}_t = \nabla h(\mathbf{x})|_{\mathbf{x}=\hat{\mathbf{x}}_{t|t-1}}$, where ∇ is the first order derivative operator with respect to $\mathbf{x}$. Then, $\tilde{\mathbf{y}}_t = \mathbf{z}_t - h(\hat{\mathbf{x}}_{t|t-1})$ is the innovation or measurement residual, $\mathbf{S}_t = \mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T + \mathbf{R}_t$ is the innovation (or residual) covariance, $\mathbf{L}_t^* = \mathbf{P}_{t|t-1}\mathbf{H}_t^T\mathbf{S}_t^{-1}$ is the size $4 \times N$ near-optimal Kalman gain which minimizes the trace of the a posteriori estimate covariance matrix $\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{L}_t^*\mathbf{H}_k)\mathbf{P}_{t|t-1}$ and finally $\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{L}_t^*\tilde{\mathbf{y}}_t$ is the updated state estimate. Note that if a particular column of $\mathbf{L}_t^*$ is a size 4×1 zero vector, i.e., $\mathbf{L}_{t,i} = \mathbf{0}$, the innovation from sensor i becomes zero. In other words, sensor i is not selected at time t.

III. Sparsity-Promoting Kalman Filtering Problem

In this section we formulate the optimal sparse Kalman filtering problem and use ADMM to solve it. Our treatment closely follows [9], [10], where ADMM was used for the identification of optimal sparse state-feedback gains. In this letter, we obtain the Kalman gain matrix, $\mathbf{L}_t$, as the solution to the optimization problem,

$$\min_{\mathbf{L}_t} \cdot J(\mathbf{L}_t) + \gamma g(\mathbf{L}_t). \tag{5}$$

In (5), $J(\mathbf{L}_t) \triangleq (1/2)\,\mathrm{tr}\{\mathbf{P}_{t|t}\}$ and $g(\mathbf{L}_t)$ is a sparsity-promoting penalty function defined as,

$$g(\mathbf{L}_t) \triangleq \mathrm{card}\left(\left[\, \|\mathbf{L}_{t,1}\|_2 \;\; \|\mathbf{L}_{t,2}\|_2 \cdots \|\mathbf{L}_{t,N}\|_2 \,\right]\right) \tag{6}$$

where $\mathbf{L}_{t,i}$ is the i^{th} column of $\mathbf{L}_t$, $\|.\|_2$ denotes the 2-norm, and $\mathrm{card}(.)$ denotes the cardinality function, i.e., the number of nonzero elements of a vector. $g(\mathbf{L}_t)$ then represents the number of non-zero columns of a matrix. Our goal is to promote sparsity of the Kalman gain by incorporating the $g(\mathbf{L}_t)$ function into the optimization problem. The emphasis on the sparsity of the columns of $\mathbf{L}_t$ is represented by the scalar $\gamma > 0$. The column sparsity of $\mathbf{L}_t$ translates to fewer active sensors. A larger γ encourages fewer nonzero columns in $\mathbf{L}_t$, while $\gamma = 0$ renders the Kalman gain $\mathbf{L}_t = \mathbf{L}_t^*$ found from the EKF.

Since the problem given in (5) involves the $\mathrm{card}(.)$ function, it is a combinatorial problem [17]. The alternating direction method of multipliers (ADMM) combines the decomposability of dual ascent with the convergence properties of the method of multipliers [8], [9]. By introducing the new optimization variable $\mathbf{G}_t$, problem (5) can be rewritten in the equivalent form,

$$\begin{aligned} \min_{\mathbf{L}_t, \mathbf{G}_t} \cdot \quad & J(\mathbf{L}_t) + \gamma g(\mathbf{G}_t) \\ \text{s.t.} \quad & \mathbf{L}_t - \mathbf{G}_t = 0. \end{aligned} \tag{7}$$

For the constrained problem in (7), the augmented Lagrangian can be written as [8],

$$\begin{aligned} \mathcal{L}_\rho(\mathbf{L}_t, \mathbf{G}_t, \Lambda) = {} & J(\mathbf{L}_t) + \gamma g(\mathbf{G}_t) \\ & + \mathrm{tr}\left(\Lambda^T(\mathbf{L}_t - \mathbf{G}_t)\right) + \frac{\rho}{2}\|\mathbf{L}_t - \mathbf{G}_t\|_F^2 . \end{aligned} \tag{8}$$

In (8), the matrix Λ is the dual variable or Lagrange multiplier associated with the constraint $\mathbf{L}_t - \mathbf{G}_t = 0$, $\|.\|_F$ represents the Frobenius norm, and ρ is a positive constant. The ADMM algorithm finds the minimum of (8) by solving the following optimization problems iteratively,

$$\mathbf{L}_t^{k+1} = \arg\min_{\mathbf{L}_t}.\ \mathcal{L}_\rho(\mathbf{L}_t, \mathbf{G}_t^k, \Lambda^k) \tag{9}$$

$$\mathbf{G}_t^{k+1} = \arg\min_{\mathbf{G}_t}.\ \mathcal{L}_\rho(\mathbf{L}_t^{k+1}, \mathbf{G}_t^k, \Lambda^k) \tag{10}$$

$$\Lambda^{k+1} = \Lambda^k + \rho\left(\mathbf{L}_t^{k+1} - \mathbf{G}_t^{k+1}\right) \tag{11}$$

until both of the following convergence criteria are met,

$$\left\|\mathbf{L}_t^{k+1} - \mathbf{G}_t^{k+1}\right\|_F \le \epsilon_1 \quad \text{and} \quad \left\|\mathbf{G}_t^{k+1} - \mathbf{G}_t^{k}\right\|_F \le \epsilon_2.$$

In the above equations t denotes the time step of tracking and k denotes the ADMM iteration. We henceforth refer to solving problems (9) and (10) as the L-minimization and G-minimization steps, respectively. The ADMM algorithm iteratively solves for $\mathbf{L}_t$ when $\mathbf{G}_t$ is held fixed, solves for $\mathbf{G}_t$ when $\mathbf{L}_t$ is held fixed, and then updates the dual variable. These iterations are repeated until convergence is achieved. For nonconvex optimization problems, a proof of convergence for the ADMM algorithm can not be established. However, in practice it has been observed that ADMM works well when the value of ρ is chosen to be large [9]. This can be attributed to the fact that for large ρ the term corresponding to the Frobenius norm dominates the augmented Lagrangian.

A. L-*Minimization Step*

As demonstrated in [8], [9], the L-minimization step in (9) can be rewritten as,

$$\min_{\mathbf{L}_t} \cdot \varphi(\mathbf{L}_t) = J(\mathbf{L}_t) + \frac{\rho}{2}\|\mathbf{L}_t - \mathbf{U}^k\|_F^2 \tag{12}$$

where $\mathbf{U}^k \triangleq \mathbf{G}_t^k - (1/\rho)\Lambda^k$. Let $J(\mathbf{L}_t) \triangleq (1/2)\,\mathrm{tr}(\mathbf{P}_{t|t})$. Using the property $\|\mathbf{L}_t - \mathbf{U}^k\|_F^2 = \mathrm{tr}\left[(\mathbf{L}_t - \mathbf{U}^k)(\mathbf{L}_t - \mathbf{U}^k)^T\right]$, the objective function can be written as,

$$\varphi(\mathbf{L}_t) = \frac{1}{2}\,\mathrm{tr}\left(\mathbf{P}_{t|t}\right) + \frac{\rho}{2}\,\mathrm{tr}\left[(\mathbf{L}_t - \mathbf{U}^k)(\mathbf{L}_t - \mathbf{U}^k)^T\right]. \tag{13}$$

The *a posteriori* estimate covariance matrix $\mathbf{P}_{t|t}$ has the form,

$$\begin{aligned}\mathbf{P}_{t|t} &= \mathrm{cov}\{\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}\} \\ &= (\mathbf{I} - \mathbf{L}_t\mathbf{H}_t)\mathbf{P}_{t|t-1}(\mathbf{I} - \mathbf{L}_t\mathbf{H}_t)^T + \mathbf{L}_t\mathbf{R}\mathbf{L}_t^T.\end{aligned} \tag{14}$$

Taking the derivative of $\varphi(\mathbf{L}_t)$ with respect to $\mathbf{L}_t$ and setting it to zero, yields the optimal $\mathbf{L}_t^{k+1}$ for the minimization problem (12),

$$\begin{aligned}\nabla\varphi(\mathbf{L}_t) &= \nabla\left\{\frac{1}{2}\,\mathrm{tr}\{(\mathbf{I}-\mathbf{L}_t\mathbf{H}_t)\mathbf{P}_{t|t-1}(\mathbf{I}-\mathbf{L}_t\mathbf{H}_t)^T + \mathbf{L}_t\mathbf{R}\mathbf{L}_t^T\}\right. \\ &\quad \left. + \frac{\rho}{2}\,\mathrm{tr}\left[(\mathbf{L}_t - \mathbf{U}^k)(\mathbf{L}_t - \mathbf{U}^k)^T\right]\right\} \\ &= -\left(\mathbf{H}_t\mathbf{P}_{t|t-1}\right)^T + \mathbf{L}_t\mathbf{S}_t + \rho(\mathbf{L}_t - \mathbf{U}^k) = 0.\end{aligned}$$

Hence,

$$\mathbf{L}_t^{k+1} = \left[(\mathbf{H}_t\mathbf{P}_{t|t-1})^T + \rho\mathbf{U}^k\right](\mathbf{S}_t + \rho\mathbf{I})^{-1} \tag{15}$$

where $\mathbf{I}$ is the identity matrix of size $N \times N$. Note that, $\rho = 0$ yields the Kalman gain $\mathbf{L}_t^*$.

B. G-*Minimization Step*

For the G-minimization step, we generalize the proof in [9] to the case of column-partitioned matrices. Similar to the L-minimization step in (12), the G-minimization step in (10) can be rewritten as,

$$\min_{\mathbf{G}_t} \cdot \phi(\mathbf{G}_t) = \gamma g(\mathbf{G}_t) + \frac{\rho}{2}\left\|\mathbf{G}_t - \mathbf{V}^k\right\|_F^2 \tag{16}$$

where $\mathbf{V}^k \triangleq (1/\rho)\Lambda^k + \mathbf{L}^{k+1}$. Using the separability properties of both the cardinality function and the Frobenius norm, we can rewrite the expression for ϕ as,

$$\begin{aligned}\phi(\mathbf{G}_t) &= \gamma\,\mathrm{card}\left(\left[\,\|\mathbf{G}_{t,1}\|_2\ \|\mathbf{G}_{t,2}\|_2\ \cdots \|\mathbf{G}_{t,N}\|_2\right]\right) \\ &\quad + \frac{\rho}{2}\left\|\mathbf{G}_t - \mathbf{V}^k\right\|_F^2 \\ &= \sum_{i=1}^{N}\Psi(\mathbf{G}_{t,i})\end{aligned} \tag{17}$$

where $\mathbf{G}_{t,i}$ is the i^{th} column of $\mathbf{G}_t$, $\Psi(\mathbf{G}_{t,i}) \triangleq \gamma\,\mathrm{card}\left(\|\mathbf{G}_{t,i}\|_2\right) + (\rho/2)\left\|\mathbf{G}_{t,i} - \mathbf{V}_i^k\right\|_2^2$ and $\mathbf{V}_i^k$ is the i^{th} column of $\mathbf{V}^k$. Therefore,

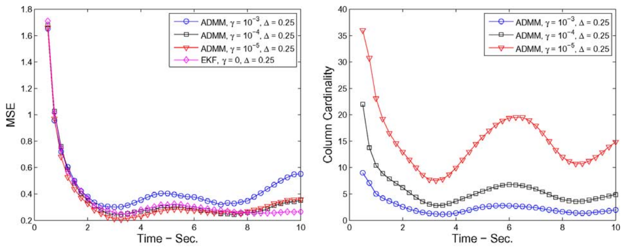
$$\Psi(\mathbf{G}_{t,i}) = \begin{cases}\gamma + \frac{\rho}{2}\left\|\mathbf{G}_{t,i} - \mathbf{V}_i^k\right\|_2^2, & \|\mathbf{G}_{t,i}\|_2 \ne 0 \\ \frac{\rho}{2}\left\|\mathbf{V}_i^k\right\|_2^2, & \|\mathbf{G}_{t,i}\|_2 = 0\end{cases}. \tag{18}$$

Note that, for $\|\mathbf{G}_{t,1}\|_2 \ne 0$, $\mathbf{G}_{t,i} = \mathbf{V}_i^k$ yields the minimum of $\Psi(\mathbf{G}_{t,i})$ as $\Psi(\mathbf{G}_{t,i}) = \gamma$. Otherwise, $\|\mathbf{G}_{t,1}\|_2 = 0$ yields the minimum of $\Psi(\mathbf{G}_{t,i})$ as $\Psi(\mathbf{G}_{t,i}) = (\rho/2)\left\|\mathbf{V}_i^k\right\|_2^2$. In conclusion, $\mathbf{G}_{t,i}^{k+1}$ minimizes $\Psi(\mathbf{G}_{t,i})$ according to,

$$\mathbf{G}_{t,i}^{k+1} = \begin{cases}0 & \frac{\rho}{2}\left\|\mathbf{V}_i^k\right\|_2^2 \le \gamma \\ \mathbf{V}_i^k & \frac{\rho}{2}\|\mathbf{V}_i^k\|_2^2 > \gamma\end{cases}. \tag{19}$$

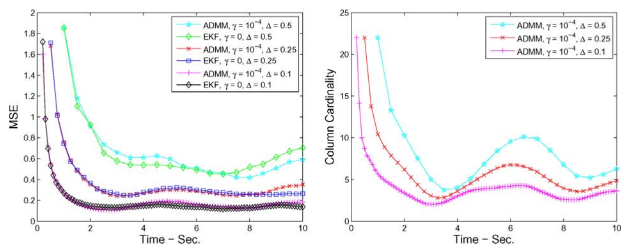
IV. Simulation Results

In this section, we illustrate the utility of the developed approach with simulation results. We consider an ROI of size $b^2 = 50 \times 50\ \mathrm{m}^2$, and $N = 36$ sensors that are uniformly deployed. The target emits energy $P_0 = 1000$, and process noise parameter is selected as $\tau = 10^{-2}$. The initial point of each target $\mathbf{x}_0$ is generated from the probability density function $p(\mathbf{x}_0) \sim \mathcal{N}(\boldsymbol{\mu}_0, \mathcal{P}_0)$, which is assumed to be Gaussian with mean $\boldsymbol{\mu}_0 = [-20\ \ -20\ \ 2\ \ 2]^T$ with the covariance matrix $\mathcal{P}_0 = \mathrm{diag}[2.778, 2.778, 0.01, 0.01]$. Note that the elements inside the $\mathrm{diag}[.]$ function are located on the main diagonal of $\mathcal{P}_0$. Then, $\mathbf{x}_0$ remains in the ROI with high probability. The target is observed for $T_d = 10$ seconds. In our simulations, the sampling interval Δ is varied and the target location is estimated between time steps $[2\Delta, T_d]$. Let $\mathrm{MSE}(t)$ be the MSE at time



(a) (b)

Fig. 1. (a) Tracking performance for different values of γ, $\Delta = 0.25$, $\rho = 2$ (b) Column cardinality for different values of γ, $\Delta = 0.25$, $\rho = 2$.



(a) (b)

Fig. 2. (a) Tracking performance under different values of Δ, $\gamma = 10^{-4}$, $\rho = 2$ (b) Column cardinality under different values of Δ, $\gamma = 10^{-4}$, $\rho = 2$.

step t which is averaged over $T_{trials} = 1000$ trials and defined as,

$$\mathrm{MSE}(t) \triangleq \frac{1}{T_{trials}} \sum_{c=1}^{T_{trials}} (x_c(t) - \hat{x}_c(t))^2 + (y_c(t) - \hat{y}_c(t))^2 \tag{20}$$

where $t \in \{2\Delta, 3\Delta, \ldots, T_d\}$, $(x_c(t),\, y_c(t))$ is the target location for c^{th} trial, at time step t and $(\hat{x}_c(t),\, \hat{y}_c(t))$ are their corresponding estimates. The ADMM parameters are selected as $\epsilon_1 = \epsilon_2 = 10^{-3}$ and $\rho = 2$. The ADMM iterations are initialized with $\mathbf{G}_t = \Lambda_t = \mathbf{0}$ and $\mathbf{L}_t$ is found from (15). Simulation results show that the required number of ADMM iterations until meeting the stopping criterion is around 10. In general, the convergence of ADMM for a nonconvex problem, such as the one considered in our work, is not guaranteed [8]. When it does converge, the final result can depend on the choice of ρ and the initial values of $\mathbf{G}_t$, Λ_t and $\mathbf{L}_t$ [8].

In Fig. 1, we compare the estimation performance of ADMM with standard EKF which corresponds to $\gamma = 0$ in (5) and all N sensor transmit their measurements. The ADMM algorithm is executed for different values of γ, i.e, $\gamma = 10^{-3}$, $\gamma = 10^{-4}$ and $\gamma = 10^{-5}$. Simulation results show that as the sparsity promoting parameter γ increases, $\mathbf{L}_t$ becomes sparser and the number of sensors selected at each iteration decreases. Note that, using $\gamma = 10^{-4}$, around 5 informative sensors are selected out of $N = 36$ sensors and the estimation performance is still very close to that of the EKF case. In Fig. 2, we fix the sparsity promoting parameter to $\gamma = 10^{-4}$ and vary the sampling interval Δ. Simulation results show that decreasing the sampling interval not only decreases the estimation error in tracking but also provides a sparser Kalman gain matrix.

V. Conclusion

In this letter, we have addressed the problem of target tracking based on received signal strengths. Rather than using an EKF, which uses all the sensor measurements, the Kalman gain matrix has been obtained as the solution to an optimization problem in which a sparsity-promoting penalty function is added to the objective. Since each column of the Kalman gain matrix corresponds to one sensor measurement, by formulating an optimization problem in which the number of nonzero columns of the Kalman gain are penalized, we promote the use of fewer sensors by the Kalman filter. Simulation results show that using a subset of sensors, which are obtained from the non-zero columns of the Kalman gain matrix, can achieve tracking performance that is very close to that of EKF. The usage of sparsity in other sensor management problems, such as the resource allocation or sensor deployment problems, will be investigated as future research directions.

References

[1] X. Sheng and Y.-H. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 44–53, Jan. 2005.

[2] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, Feb. 2009.

[3] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 61–72, Mar. 2002.

[4] G. Hoffmann and C. Tomlin, "Mobile sensor network control using mutual information methods and particle filters," *IEEE Trans. Automat. Contr.*, vol. 55, no. 1, pp. 32–47, Jan. 2010.

[5] E. Masazade, R. Niu, P. K. Varshney, and M. Keskinoz, "Energy aware iterative source localization for wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4824–4835, Sep. 2010.

[6] L. Zuo, R. Niu, and P. K. Varshney, "Posterior CRLB based sensor selection for target tracking in sensor networks," in *IEEE Int. Conf. Acoust., Speech and Signal Process.*, Apr. 2007, vol. 2, pp. II-1041–II-1044.

[7] A. Carmi, "Sensor scheduling via compressed sensing," in *13th Conf. Inform. Fusion*, Jul. 2010, pp. 1–8.

[8] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[9] F. Lin, M. Fardad, and M. Jovanovic, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Trans. Automatic. Contr.*, 2011 [Online]. Available: arXiv:1111.6188, submitted for publication

[10] F. Lin, M. Fardad, and M. R. Jovanović, "Sparse feedback synthesis via the alternating direction method of multipliers," in *Proc. 2012 Amer. Contr. Conf.*.

[11] M. Fardad, F. Lin, and M. R. Jovanović, "On the optimal synchronization of oscillator networks via sparse interconnection graphs," in *Proc. 2012 Amer. Contr. Conf.*.

[12] F. Lin, M. Fardad, and M. R. Jovanović, "Algorithms for leader selection in large dynamical networks: Noise-corrupted leaders," in *Proc. 50th IEEE Conf. Dec. and Contr.*, 2011, pp. 2932–2937.

[13] M. Fardad, F. Lin, and M. R. Jovanović, "Sparsity-promoting optimal control for a class of distributed systems," in *Proc. 2011 Amer. Contr. Conf.*, 2011, pp. 2050–2055.

[14] S. Farahmand, G. Giannakis, G. Leus, and Z. Tian, "Sparsity-aware Kalman tracking of target signal strengths on a grid," in *Proc. 14th Int. Conf. Inform. Fusion*, Jul. 2011, pp. 1–6.

[15] R. Niu and P. K. Varshney, "Distributed detection and fusion in a large wireless sensor network of random size," *EURASIP J. Wireless Commun. Network.*, vol. 2005, no. 4, pp. 462–472, 2005.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall, 1995.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ Press, 2004.